

12

DEMANDE DE BREVET D'INVENTION

A1

22 Date de dépôt : 15.09.98.

30 Priorité :

43 Date de mise à la disposition du public de la demande : 17.03.00 Bulletin 00/11.

59 Liste des documents cités dans le rapport de recherche préliminaire : *Se reporter à la fin du présent fascicule*

60 Références à d'autres documents nationaux apparentés :

71 Demandeur(s) : THOMSON MULTIMEDIA Société anonyme — FR.

72 Inventeur(s) : GUILLOTEL PHILIPPE.

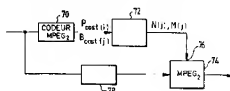
73 Titulaire(s) :

74 Mandataire(s) : THOMSON MULTIMEDIA.

54 PROCÉDE DE COMPRESSION D'IMAGES ET DISPOSITIF POUR LA MISE EN ŒUVRE DE CE PROCÉDE.

57 L'invention est relative à un procédé de compression d'images, notamment du type MPEG2, dans lequel les images sont codées suivant des groupes (GOP) dont chacun comprend N images, avec une image I codée en intra, des images P prédites en fonction de l'image intra I ou d'une image P précédente, chaque image P étant précédée et suivie par n images B prédites bidirectionnellement, n pouvant être nul. Le nombre M = n + 1 représente la structure du groupe.

On détermine à l'aide d'un codage d'essai (70) au moins un paramètre caractérisant les images sources devant être codées suivant un groupe et on fait dépendre les nombres N et M de ce(s) paramètre(s). Au cours du codage d'essai, on confère des valeurs déterminées à N, M et au pas de quantification Q.



PROCEDE DE COMPRESSION D'IMAGES ET DISPOSITIF POUR LA MISE EN
OEUVRE DE CE PROCEDE

L'invention est relative à un procédé de compression d'images dans lequel les images sont codées suivant des groupes de longueurs variables.

Elle concerne plus particulièrement un procédé de type
5 MPEG, notamment de type MPEG2. Bien que l'invention ne soit pas limitée à cette norme, dans la suite de la description, on se référera principalement à celle-ci.

On rappelle ci-après le principe d'une telle compression.

10 Dans la norme MPEG2 vidéo, on obtient une compression des signaux numériques vidéo en exploitant la redondance spatiale et la redondance temporelle des images codées.

La redondance spatiale est évaluée principalement grâce à la succession de trois opérations : une opération communément appelée transformée cosinus discrète et notée DCT
15 ("Discrete Cosine Transform"), une opération de quantification des coefficients issus de la DCT et une opération de codage à longueur variable pour décrire les coefficients quantifiés issus de la DCT.

20 La redondance temporelle est analysée par une opération de compensation de mouvement qui consiste, par translation

de chaque bloc de l'image courante, à rechercher le bloc le plus ressemblant situé dans une image de référence. L'analyse de la redondance temporelle conduit à déterminer un champ de vecteurs de translation, communément appelés vecteurs de mouvement, ainsi
5 qu'une erreur de prédiction qui est la différence entre le signal de l'image courante et le signal de l'image prédite par compensation de mouvement. L'erreur de prédiction est alors analysée selon le principe de la redondance spatiale.

Le codage MPEG2 est de type prédictif. Il s'ensuit que
10 le décodage qui lui est associé doit être régulièrement réinitialisé afin de protéger le signal contre toute erreur de transmission ou toute rupture de signal due au basculement du décodeur d'un programme sur un autre.

A cette fin, la norme MPEG2 prévoit que, périodi-
15 quement, les images doivent être codées en mode spatial, c'est-à-dire selon un mode exploitant seulement la redondance spatiale. Les images codées en mode spatial sont appelées images INTRA ou images I.

Les images codées en exploitant la redondance tempo-
20 relle sont de deux types : d'une part, les images construites par référence à une image temporellement antérieure à partir d'une prédiction avant et, d'autre part, les images construites par référence à deux images temporellement antérieure et postérieure à partir d'une prédiction avant et d'une prédiction
25 arrière.

Les images codées construites à partir d'une prédiction avant sont appelées images prédites ou images P et les images codées construites à partir d'une prédiction avant et d'une prédiction arrière sont appelées images bidirectionnelles
30 ou images B.

Une image I est décodée sans faire référence à d'autres images qu'à elle-même. Une image P est décodée en faisant référence à l'image P ou I qui la précède. Une image B est décodée en faisant appel à l'image I ou P qui la précède et
35 à l'image I ou P qui la suit.

La périodicité des images I définit un groupe d'images communément noté GOP ("Group Of Pictures").

A l'intérieur d'un même GOP, la quantité de données contenues dans une image I est généralement supérieure à la
5 quantité de données contenues dans une image P et la quantité de données contenues dans une image P est généralement supérieure à la quantité de données contenues dans une image B.

En 50 Hertz, le GOP se présente comme une image I suivie d'une séquence d'images B et P qui présente, la plupart du
10 temps, la séquence suivante

I, B, B, P, B, B, P, B, B, P, B, B.

Cependant, la norme n'oblige pas à prévoir, comme c'est le cas en général, $N = 12$ images dans un GOP, ni à prévoir que les distances M entre deux images P soient toujours égales à
15 3. De façon plus précise, la distance M est le nombre n d'images B précédant ou suivant une image P, augmenté d'une unité, soit $M = n + 1$.

Le nombre N représente la taille ou longueur du GOP, tandis que le nombre M représente sa structure.

20 L'invention résulte de la constatation que l'on peut agir sur les paramètres M et N pour améliorer le taux de compression et/ou améliorer la qualité du codage.

Le procédé de codage selon l'invention est caractérisé en ce qu'on détermine au moins un paramètre caractérisant les
25 images sources devant être codées suivant un groupe et en ce qu'on fait dépendre la longueur et la structure du groupe de ce(s) paramètre(s).

Dans un mode de réalisation, le (les) paramètre(s) caractérisant les images sources est déterminé à l'aide d'un
30 codage d'essai au cours duquel on confère des valeurs déterminées à N, M et au pas Q de quantification.

Le codage d'essai est par exemple réalisé en boucle ouverte.

Dans une réalisation particulièrement simple, on
35 détermine séparément un paramètre (Pcost) caractérisant les

images P obtenues lors du codage d'essai et un paramètre (Bcost) caractérisant les images B obtenues lors du codage d'essai, ces paramètres caractérisant les images P et B étant, de préférence, les coûts moyens de codage des images P et B. Le coût de codage
 5 d'une image est le nombre de bits (en-têtes compris) nécessaire pour le codage.

Dans ce cas, on peut faire dépendre le nombre N du paramètre caractérisant les images P et le nombre M du paramètre caractérisant les images B.

10 Lors d'essais effectués dans le cadre de l'invention, sur des séquences d'images de natures diverses, on a constaté qu'il existait, pour chaque nature de séquence, un nombre N optimal fournissant un minimum de coût de codage (ou débit) pour les images P et un nombre M optimal fournissant un minimum de
 15 coût de codage (ou débit) pour les images B, ces coûts étant obtenus lors du codage d'essai. Ces séquences se distinguent les unes des autres par des mouvements d'amplitudes variables, des objets différents, des définitions spatiales différentes et des contenus différents.

20 On a en outre constaté qu'il existe une relation pratiquement linéaire entre le nombre N optimal et le débit des images P. De même, il existe une relation pratiquement linéaire entre le nombre M et le débit des images B. Ainsi, connaissant les débits des images P et B, il est aisé de calculer les
 25 nombres N et M fournissant les meilleurs résultats.

Dans un exemple correspondant à la norme MPEG2, 50 Hz, le codage d'essai s'effectue avec $N = 12$, $M = 3$ et $Q = 15$, la relation entre N et le débit des images P est approximativement la suivante :

$$30 \quad (1) \quad N = \text{INT} \left[\frac{389000 - P \text{ cost}}{10000} \right] + 1, \text{ avec } 12 \leq N \leq 30,$$

et la relation entre M et le débit, ou coût, Bcost des images B est la suivante :

$$(2) \quad M = \text{INT} \left[\frac{179000 - B \text{ cost}}{20000} \right] + 1, \text{ avec } 1 \leq M \leq 7.$$

Il est également possible de limiter M à 5.

Dans ces formules, INT signifie la partie entière.

La limitation de N entre 12 et 30 et la limitation de M à une valeur maximale de 7 permet une réalisation simple des
 5 codeurs et de limiter le temps de changement de programme. Dans le même but, on peut aussi imposer d'autres limitations ou contraintes, notamment que M soit constant dans le GOP et/ou qu'il soit un sous-multiple de N.

Dans une réalisation, si les valeurs de N et de M
 10 prises individuellement et ensemble ne sont pas compatibles avec les contraintes, on choisira des valeurs de M et de N les plus proches des valeurs calculées et qui satisfont à la compatibilité imposée. Dans ce cas, on privilégiera la valeur de M, c'est-à-dire que si un choix doit être effectué entre plusieurs
 15 couples M, N, on choisira le couple pour lequel la valeur M est la plus proche de celle qui résulte du calcul.

La formule (2) ci-dessus ne s'applique que si Bcost ne dépasse pas 179000. Dans le cas contraire, c'est-à-dire si Bcost > 179000, l'expérience a montré qu'il fallait, dans cet
 20 exemple, que M soit choisi de la façon suivante :

$$(3) M = 5 \cdot \text{INT} \left[\frac{P \text{ cost}}{B \text{ cost}} - 1 \right], \text{ avec } 1 \leq M \leq 7.$$

Si le coût d'une image B est supérieur au coût d'une image P, il est préférable que le GOP ne contienne pas d'image B, c'est-à-dire M = 1. En effet, les images P présentant une
 25 meilleure qualité de prédiction que les images B et étant, par hypothèse, de coût inférieur, la présence de telles images B constituerait un inconvénient dans ce cas.

Les coûts, en bits, de chaque image P et de chaque image B sont déterminés, par exemple, au fur et à mesure de
 30 l'apparition de ces images. Dans une réalisation, les valeurs de M et N sont sélectionnées en effectuant une moyenne sur toutes les images P et B du codage d'essai, le codage proprement dit n'étant effectué qu'après le codage d'essai de N images sources,

N étant déterminé par le coût de codage des images P. Dans ce cas, le paramètre M peut rester constant dans le GOP.

Dans une autre réalisation, qui permet une adaptation plus rapide aux variations de contenus des scènes ainsi qu'une
5 réduction du retard entre l'arrivée des images sources et le codage proprement dit (et qui permet donc une mémoire tampon de capacité plus réduite), on fait démarrer le codage proprement dit dès que le codage d'essai fournit des données permettant ce démarrage. Ainsi, la première image B du codage d'essai fournit
10 un nombre M permettant de démarrer le codage et le nombre N est fourni par la première image P du codage d'essai. Il est également possible de ne faire débiter le codage qu'après le codage d'essai de la première image P ; dans ce cas, le codage démarre lorsqu'une valeur de N et une valeur de M sont connues.

15 Avec ce type de codage "à la volée", le nombre M, c'est-à-dire la structure, peut varier à l'intérieur d'un GOP, ce qui permet une adaptation rapide aux variations de contenu de la scène.

Dans ce codage effectué au fur et à mesure, le GOP est
20 interrompu quand le nombre d'images déjà codées dans le GOP courant est au moins égal au nombre N mesuré (par Poost dans l'exemple ci-dessus), ou lors d'un changement de scène.

Pour éviter des variations importantes des paramètres entre groupes qui se succèdent, il peut s'avérer utile de
25 s'écarter des valeurs calculées. Par exemple, si le calcul montre que, pour une grande partie de la longueur du GOP, par exemple au moins 80%, il faudrait $M = 1$, alors que, pour le reste du GOP, le calcul montre que M doit être supérieur à 1, on adopte, malgré tout, la valeur 1 pour M, même si le calcul
30 montre qu'il faut une valeur différente.

De façon similaire, si pour le GOP précédent, $M = 1$ et si, pour le GOP courant, le calcul montre qu'il faudrait une valeur $M = 1$ pour une partie importante du GOP courant, par exemple au moins 60%, on adopte aussi la valeur 1 pour M, même

si le résultat du calcul, tel qu'il résulte de la formule (2) ci-dessus, implique une valeur différente.

On sait que lorsqu'un changement de scène intervient, c'est-à-dire lorsqu'il apparaît une discontinuité dans la suite
5 des images vidéo, il est nécessaire d'adapter les groupes d'images GOP de part et d'autre de la discontinuité de façon telle que le nouveau groupe, qui démarre avec une image I, corresponde à la nouvelle scène.

Dans un mode de réalisation, si le changement de scène
10 intervient dans un groupe, la scène nouvelle constitue l'image I d'un groupe nouveau, le groupe altéré étant raccourci pour s'arrêter avant cette scène nouvelle si le changement de scène intervient, dans le groupe altéré, à une distance du début au moins égale au nombre minimum admissible pour N. Le début du
15 groupe altéré est utilisé pour allonger le groupe qui le précède lorsque la somme du nombre d'images précédant le changement de scène dans le groupe altéré et du nombre d'images du groupe qui le précède ne dépasse pas le maximum admissible pour N. Dans ce groupe précédent ainsi modifié (raccourci ou allongé), il peut
20 être nécessaire de modifier le nombre M précédemment calculé pour ce GOP.

Dans une variante, qui s'utilise de préférence dans le cas où la longueur du groupe altéré est inférieure au nombre admissible pour N, quand un changement de scène intervient dans
25 un groupe, la scène nouvelle constitue l'image I d'un groupe nouveau, ce groupe nouveau ayant une longueur égale à la moyenne de la longueur du groupe avant altération et de la longueur du groupe qui le précède. Avec cette variante, il peut être nécessaire de modifier le nombre M précédemment calculé pour les GOP.

30 Quand deux modifications sont possibles, par exemple quand la longueur du groupe altéré est inférieure au minimum admissible pour N, on peut effectuer un choix entre ces deux modifications en effectuant, pour chaque modification, un calcul de la distance du couple (M,N) obtenu ou couple M,N avant modi-

fication et en sélectionnant le couple pour lequel la distance est la plus faible.

Pour déterminer les paramètres N et M, on peut faire appel à la mesure d'autres paramètres que la mesure des débits.

- 5 Par exemple, on peut faire appel, pour déterminer N, à l'énergie des images Intra I. On peut également déterminer l'amplitude des mouvements ou l'erreur de compensation de mouvement dite DFD (Displaced Frame Difference) pour la détermination de M et N.

- D'autres caractéristiques et avantages de l'invention
10 apparaîtront avec la description de certains de ses modes de réalisation, celle-ci étant effectuée en se référant aux dessins ci-annexés, sur lesquels :

la figure 1 illustre un macrobloc au standard 4.2.0,

la figure 2 est un schéma illustrant la transformation

- 15 DCT,

la figure 3 montre un groupe d'images GOP selon la norme MPEG ou analogue,

les figures 4 à 7 sont des diagrammes illustrant le procédé conforme à l'invention, et

- 20 la figure 8 est un schéma d'un montage pour la mise en oeuvre du procédé selon l'invention.

On se réfère tout d'abord aux figures 1 à 3 qui visent à rappeler certains principes utilisés dans le codage MPEG2.

- Dans la norme MPEG2, on peut partir d'une image com-
25 portant, en mode progressif, 576 lignes de 720 points chacune. En mode entrelacé, cette image est composée de deux trames dont chacune comprend 288 lignes, également de 720 points chacune.

- Chaque image est décomposée en macroblocs dont chacun est formé par un carré de 16X16 points de luminance. Chaque
30 macrobloc est ainsi formé de quatre blocs carrés de 8X8 points de luminance. A chacun de ces quatre blocs de luminance, on associe (dans le format 4.2.0) deux blocs de chrominance dont chacun présente 8X8 points, l'un des blocs représentant le
35 l'autre bloc représentant le signal Cb de différence de couleur

ou chrominance bleue. Dans le format 4.2.2, à chaque macrobloc de luminance, on associe quatre blocs de chrominance 8x8, deux blocs pour la chrominance bleue et deux blocs pour la chrominance rouge. Il existe aussi un format 4.4.4 pour lequel
5 chacune des composantes de luminance et de chrominances comporte chacune quatre blocs 8x8.

On a représenté sur la figure 1 quatre blocs de luminance 8x8, de référence 10 dans l'ensemble et des blocs 12 et 14 de chrominance 8x8 pour les chrominances respectivement bleue et
10 rouge, l'ensemble illustrant un macrobloc au standard 4.2.0.

Chaque bloc est codé en utilisant une transformation notée DCT qui est une transformation cosinus discrète qui permet de transformer un bloc de luminance (par exemple) en un bloc de coefficients représentant des fréquences spatiales. Comme on
15 peut le voir sur la figure 2, un bloc source 16 est transformé en un bloc 18 de 8x8 coefficients. Le coin supérieur gauche 20 du bloc 18 correspond aux fréquences spatiales nulles (valeur moyenne du bloc) et, à partir de cette origine 20, les fréquences horizontales augmentent vers la droite, comme représenté
20 par la flèche 22, tandis que les fréquences spatiales verticales augmentent en partant du haut vers le bas, comme représenté par la flèche 24.

Pour chaque macrobloc, il faut choisir le type de codage : soit "intra", soit "inter". Un codage intra consiste à
25 appliquer la transformation DCT sur un bloc source de l'image, tandis qu'un codage inter consiste à appliquer la transformation DCT sur un bloc représentant la différence entre un bloc source et un bloc prédit, ou bloc de prédiction, d'une image précédente ou suivante.

30 Le choix dépend en partie du type d'images auquel appartient le macrobloc. Ces images sont de trois types : le premier type est le type dit I, ou intra, pour lequel le codage est intra pour tous les macroblocs.

Le second type est de type P ou de prédiction ; dans
35 ce type d'images, le codage de chaque macrobloc peut être soit

intra soit inter. Dans le cas d'un codage inter sur une image de type P, la transformation DCT est appliquée sur la différence entre le macrobloc courant de cette image P et un macrobloc de prédiction issu de l'image I ou P précédente.

- 5 Le troisième type d'images est appelé B ou bidirectionnel. Chaque macrobloc d'un tel type d'image est soit codé en intra, soit codé en inter. Le codage en inter consiste également à appliquer la transformation sur la différence entre le macrobloc courant de cette image B et un macrobloc de prédiction.
- 10 Ce macrobloc de prédiction peut être issu soit de l'image précédente soit de l'image suivante ou des deux à la fois (prédiction bidirectionnelle), les images de prédiction dites précédente ou suivante ne pouvant être que de type I ou P.

- On a représenté sur la figure 3 un ensemble d'images
- 15 formant un groupe appelé GOP ("Group Of Pictures") qui comprend douze images, à savoir une image I suivie de onze images B et P selon la succession suivante : B, B, P, B, B, P, B, B, P, B, B.

- Un GOP est caractérisé par une longueur, c'est-à-dire un nombre d'images N qui, dans un exemple, ne peut être compris
- 20 que entre 12 et 30, et par un paramètre de structure M représentant la distance entre deux images P, c'est-à-dire le nombre d'images B entre deux images P successives, augmenté d'une unité. Dans cet exemple, ce paramètre M est égal à 3. Egalement à titre d'exemple, ce nombre M peut être compris entre 1 (pas
- 25 d'image B) et 7. En outre, on impose que ce nombre M soit un sous multiple du nombre N pour simplifier le codeur.

Jusqu'à présent, les images sont codées en gardant N et M constants dans le codeur.

- L'invention résulte de la constatation qu'il existe
- 30 des valeurs optimales de M et de N qui diffèrent selon des séquences d'images codées. En effet, selon que les séquences d'images présentent une définition plus ou moins importante et un mouvement plus ou moins important, les valeurs optimales de M et de N peuvent différer de façon importante. Par valeur opti-

male, on entend celles qui, pour une même qualité, exigent un nombre minimum de bits.

Au surplus, les études expérimentales effectuées dans le cadre de l'invention ont montré que la taille optimale N_{opt} du GOP pour une séquence déterminée d'images correspond à la valeur minimale P_{cost} , sur cette séquence, du nombre de bits qu'il faut utiliser pour coder (en-têtes compris) les images P. Cette propriété est illustrée par le diagramme de la figure 4 sur lequel on a porté, en abscisses, le nombre N et, en ordonnées, la valeur P_{cost} pour une séquence notée i . Cette valeur P_{cost} est le nombre de bits à utiliser pour coder une image P en valeur moyenne sur la séquence i . On voit, ainsi, que la valeur $P_{cost}(i)$ est représentée par une courbe 32 présentant un minimum 34 pour lequel la valeur de N est optimale (N_{opt}).

De façon analogue, on a constaté que la valeur optimale du nombre M correspond au minimum $B_{cost}(i)$ du nombre de bits à utiliser en valeur moyenne pour coder des images B sur une séquence déterminée, notée i . Ainsi, sur le diagramme de la figure 5, le nombre M a été représenté en abscisses et, en ordonnées, on a porté le nombre $B_{cost}(i)$. On voit, sur ce diagramme, que la courbe 36 présente un minimum 38 correspondant à la valeur optimale de M (M_{opt}).

Des mesures ont été effectuées, notamment sur des séquences test classiques en codage MPEG et notées "Horse", "Flower garden" et "Mobcal". La séquence "Horse" correspond à des mouvements rapides et avec une bonne définition, la séquence "Flower garden" correspond également à une bonne définition et des mouvements moyens, tandis que la séquence "Mobcal" correspond à peu de mouvement et à une définition élevée. D'autres séquences ont également été testées, telles qu'une séquence de kayak avec des mouvements rapides et peu de définition, une séquence de basket et une séquence avec des mouvements moyens, homogènes et des images de bonne définition.

On a aussi constaté que, si le groupe est soumis à un codage d'essai avec des valeurs déterminées M , de N , et du pas Q

de quantification, ces valeurs ne correspondant pas obligatoirement aux valeurs optimales pour la séquence i considérée, le coût moyen de codage P_{cost} des images P et le coût moyen de codage B_{cost} des images B représentent respectivement N et M . En outre, comme le montre la figure 6, il existe une relation simple entre les nombres N_{opt} pour chaque séquence i et le coût de codage P_{cost} à M , N et Q déterminés. Cette relation est linéaire ou sensiblement linéaire, étant représentée par une droite 40 (figure 6) sur laquelle des points différents 42, 44, etc. représentent des séquences différentes.

La figure 7 est un diagramme sur lequel les nombres M_{opt} sont portés en abscisses et le coût de codage B_{cost} (à M , N et Q déterminés) est porté en ordonnées ; chaque point 52, 54, 56, etc. correspond à une séquence donnée. On voit que ces points se trouvent sur une droite 60. Il y a donc une relation linéaire entre M_{opt} et le coût du codage d'essai.

Lorsque les valeurs de M , N et Q , utilisées au cours du codage d'essai, sont les suivantes :

$M = 12,$
 $N = 3,$ et
 $Q = 15,$

les valeurs de M et N satisfont aux relations suivantes :

$$(1) N = \text{INT} \left[\frac{389000 - P_{\text{cost}}}{10000} \right] + 1, \text{ avec } 12 \leq N \leq 30.$$

$$(2) M = \text{INT} \left[\frac{179000 - B_{\text{cost}}}{20000} \right] + 1, \text{ avec } 1 \leq M \leq 7.$$

Bien que pour la formule (2) ci-dessus, on ait indiqué que M devait être compris entre 1 et 7, on voit sur le diagramme de la figure 7 qu'en fait, M peut être limité à 5.

Un montage destiné à la mise en oeuvre de l'invention est représenté sur la figure 8. Il comprend un premier codeur MPEG2 70 destiné à effectuer le codage d'essai ou "première passe" de codage. Ce codeur d'essai est réglé avec les paramètres fixés indiqués ci-dessus, à savoir, dans l'exemple :

M = 12, N = 3 et Q = 15. Ce codeur d'essai fonctionne, dans l'exemple, en boucle ouverte, c'est-à-dire sans régulation.

Le codeur 70 fournit les valeurs Bcost et Pcost qui sont appliquées à un convertisseur 72 qui réalise les conversions de Pcost en Nopt et Bcost en Mopt, comme représenté sur les figures 6 et 7 et conformément aux relations (1) et (2) ci-dessus.

Ces valeurs N et M sont calculées pour un groupe d'images, comme décrit ci-dessus, et sont ensuite appliquées à une entrée de commande 76 d'un codeur MPEG2 74.

Les données à l'entrée du codeur 74 sont les mêmes que celles à l'entrée du codeur d'essai 70. On prévoit donc une mémoire tampon 78 pour tenir compte du temps de traitement dans le codeur d'essai 70 et dans le convertisseur 72, cette mémoire 78 retenant les données pendant le traitement.

Dans le convertisseur 72, il est également vérifié que les couples N,M résultant des formules (1) et (2) sont compatibles avec les contraintes imposées dans le mode de réalisation, en particulier que M est un sous multiple de N. Si les valeurs résultant du calcul ne sont pas compatibles, on adopte des valeurs de N et de M les plus proches de celles qui ont été calculées, en privilégiant, cependant, la valeur M.

Le convertisseur 72 tient également compte de conditions supplémentaires.

En premier lieu, il effectue une comparaison entre Bcost et Pcost et, si Bcost est supérieur à Pcost, on affecte la valeur 1 au nombre M, le GOP ne contenant pas d'image B. En effet, dans cette hypothèse, les images B entraînent un coût de codage supérieur aux images P, il est préférable de ne conserver que des images P qui présentent une qualité de prédiction supérieure.

En second lieu, le convertisseur compare Bcost à la valeur 179000 et si Bcost dépasse 179000, la relation (2) ci-dessus est remplacée par la relation heuristique ci-dessous :

$$(3) M = 5 \cdot \text{INT} \left[\frac{P \cos t}{B \cos t} - 1 \right], \text{ avec } 1 \leq M \leq 7.$$

Le convertisseur 72 permet aussi de tenir compte de deux cas particuliers pour lesquels il est nécessaire de s'écarter de la relation (2) pour obtenir une homogénéité de qualité
5 des images.

Le premier cas est le suivant : le codage d'essai montre que M doit présenter une valeur au moins égale à 2, mais, par ailleurs, ce codage d'essai montre aussi que les valeurs
intermédiaires obtenues par M sont égales à 1 sur une grande
10 partie du groupe, par exemple au moins 80%. Dans ce cas, le convertisseur 72 impose que M soit égal à 1.

Le second cas est similaire au premier : le codage d'essai montre que M devrait être au moins égal à 2, mais les valeurs intermédiaires obtenues pour M sont égales à 1 pour au
15 moins une partie de la longueur du groupe, par exemple 60% (cette limite est inférieure à la limite prévue dans le premier cas), et le groupe précédent est tel que M = 1. Dans ce cas, on confère aussi la valeur 1 au nombre M.

Ces deux cas particuliers pour lesquels on pose la
20 valeur 1 à M résultent d'essais effectués dans le cadre de l'invention qui ont montré que ces conditions permettent, pour un même type de séquence, une bonne uniformité de qualité sur des groupes successifs.

Enfin, le convertisseur 72 tient compte des changements de scène ou "cuts" qui sont habituellement détectés dans
25 les codeurs. Quand un tel changement de scène intervient, on fait démarrer le GOP avec la nouvelle scène, c'est-à-dire que, lorsque la nouvelle scène apparaît, on lui attribue une image Intra I.

30 En outre, avec le procédé de l'invention, quand un changement de scène est détecté, on configure le GOP précédent et le GOP courant en fonction des considérations suivantes :

Si le changement de scène apparaît dans un GOP après la douzième image, le nouveau GOP démarre alors avec le changement de scène, le GOP précédent étant ainsi limité ou raccourci.

Si, par contre, le changement de scène apparaît avant la douzième image, il n'est alors plus possible de limiter le GOP précédent de façon à ce qu'il se termine juste avant le changement de scène, car, dans ce cas, son nombre d'images serait inférieur au nombre minimum imposé. On modifie alors le GOP précédent et le GOP courant de la façon suivante en distinguant deux cas.

Dans le premier cas, le changement de scène apparaît à un moment tel que la somme du nombre d'images du GOP précédent et du nombre d'images du GOP courant, juste avant le changement de scène, soit au plus égal à 30. Dans ce cas, le GOP précédent est allongé.

Dans le second cas, la somme du nombre d'images du GOP précédent et du nombre d'images du GOP courant juste avant le changement de scène est supérieur à 30. Le GOP précédent et le GOP courant sont alors réarrangés en calculant une moyenne correspondant à ces deux GOP.

Par exemple, si le GOP précédent est tel que $N = 25$ et $M = 2$ et si le changement de scène intervient après la huitième image du GOP courant pour lequel le calcul indique $N = 20$ et $M = 3$, le GOP précédent allongé du GOP courant écourté comporte 33 images. Cette valeur dépassant le maximum admissible (30), on cherche une "moyenne" correspondant à deux GOP dont le nombre total d'images est 33, chacun des GOP devant respecter les contraintes imposées. Dans ce cas, on voit qu'on peut choisir $N = 18$ et $M = 2$ pour le GOP précédent et $N = 15$ et $M = 3$ pour le GOP juste avant le changement de scène. Les longueurs 18 et 15 sont proches de la moyenne (16,5) de la longueur du groupe précédent (25) et de la longueur (8) du groupe courant altéré.

Des tests ont été réalisés sur douze séquences différentes avec des changements de scène, des éclairs lumineux (flashes) et des durées relativement importantes, et on a com-

paré les résultats obtenus avec un procédé de codage classique, correspondant à des valeurs fixes de M et de N aux résultats obtenus avec le procédé selon l'invention qui adapte les valeurs de M et de N à des séquences. Ces tests ont été réalisés avec
5 plusieurs débits. On a constaté une augmentation de qualité, mesurée par le paramètre PSNR (Peak Signal to Noise Ratio), de 0,2 dB à 1,14 dB. Cette augmentation de PSNR correspond à une économie en bits comprise entre 2 et 22% environ.

Le procédé selon l'invention peut être utilisé pour
10 tout type de procédé de compression d'images vidéo dans lequel on prévoit des images I, P et B. Il s'applique tant à l'enregistrement, en temps réel ou en temps différé, qu'à la transmission.

Le procédé n'est pas limité au cas où la taille du GOP
15 est déterminée avant codage. Il s'applique aussi quand les paramètres M et N sont calculés pour chaque image, le codage proprement dit s'effectuant à la volée. Dans ce cas, le nombre M peut varier dans un GOP, un nouveau GOP commençant par exemple quand le nombre d'images codées dans le GOP courant est au moins
20 égal au nombre N calculé. Le nombre M peut varier en fonction de la complexité des images à l'intérieur d'un GOP.

Dans ce cas, il n'est pas nécessaire de mémoriser l'ensemble du GOP dans la mémoire tampon 78 (dont la capacité peut être réduite), les contraintes sur les valeurs de M et N
25 sont réduites, étant seulement imposées par la norme MPEG2 ; les contraintes imposées lors des changements de scènes sont aussi moins sévères.

REVENDICATIONS

1. Procédé de compression d'images dans lequel les images sont codées suivant des groupes (GCP) dont chacun comprend un nombre N d'images, N représentant la longueur du groupe, lequel comporte une image I codée en intra, des images P prédites en fonction de l'image intra I ou d'une image P précédente, chaque image P étant précédée et suivie par n images B prédites bidirectionnellement, n pouvant être nul, le nombre M , qui est égal au nombre n augmenté d'une unité, représentant la structure du groupe, caractérisé en ce qu'on détermine au moins un paramètre caractérisant les images sources devant être codées suivant un groupe et en ce qu'on fait dépendre la longueur N et la structure M du groupe de ce(s) paramètre(s).
2. Procédé selon la revendication 1, caractérisé en ce que le (les) paramètre(s) caractérisant les images sources est déterminé à l'aide d'un codage d'essai au cours duquel on confère des valeurs déterminées à N , M et au pas de quantification Q .
3. Procédé selon la revendication 2, caractérisé en ce que le codage d'essai est réalisé en boucle ouverte.
4. Procédé selon la revendication 2 ou 3, caractérisé en ce que, pour caractériser les images sources, on détermine un paramètre (P_{cost}) caractérisant les images P obtenues lors du codage d'essai et un paramètre (B_{cost}) caractérisant les images B obtenues lors du codage d'essai.
5. Procédé selon la revendication 4, caractérisé en ce que le nombre N est déterminé à partir du paramètre caractérisant au moins une image P , et en ce que le nombre M est déterminé en fonction du paramètre caractérisant au moins une image B .
6. Procédé selon la revendication 4 ou 5, caractérisé en ce que les paramètres (P_{cost} , B_{cost}) caractérisant les images P et B sont les coûts de codage de ces images P et B , par exemple les coûts moyens.

7. Procédé selon les revendications 5 et 6, caracté-
risé en ce que, lorsque le coût de codage moyen (Bcost) de cha-
que image B est, au cours du codage d'essai, supérieur au coût
de codage moyen (Pcost) de chaque image P, on confère la valeur
5 1 au nombre M, le groupe ne comprenant ainsi pas d'image B.

8. Procédé selon la revendication 6 ou 7, caractérisé
en ce qu'au cours du codage d'essai, on détermine, au fur et à
mesure de l'arrivée des images sources, le coût de codage pour
chaque image B et le nombre M correspondant.

10 9. Procédé selon la revendication 8, caractérisé en ce
que, lorsque les nombres M déterminés avant la fin du codage
d'essai sont égaux à 1 pour une fraction importante du groupe,
on confère la valeur 1 au nombre M pour le groupe.

15 10. Procédé selon la revendication 8 ou 9, caractérisé
en ce que, lorsque les nombres M déterminés avant la fin du
codage d'essai sont égaux à 1 pour au moins une fraction déter-
minée du groupe, et lorsque le nombre M est égal à 1 pour le
groupe précédent, on confère la valeur 1 au nombre M pour le
groupe.

20 11. Procédé selon l'une quelconque des revendications
précédentes, caractérisé en ce qu'en cas de changement de scène
intervenant dans un groupe, la scène nouvelle constitue l'image
I d'un groupe nouveau, le groupe altéré étant raccourci pour
s'arrêter avant cette scène nouvelle si le changement de scène
25 intervient, dans le groupe altéré, à une distance du début au
moins égale au nombre minimum admissible pour N, le début du
groupe altéré étant utilisé pour allonger le groupe qui le pré-
cède lorsque la somme du nombre d'images précédant le changement
de scène dans le groupe altéré, et du nombre d'images du groupe
30 qui le précède ne dépasse pas le maximum admissible pour N.

12. Procédé selon l'une quelconque des revendications 1
à 9, caractérisé en ce qu'en cas de changement de scène inter-
venant dans un groupe, la scène nouvelle constitue l'image I
d'un groupe nouveau, le groupe altéré et le groupe qui le pré-
cède étant réarrangés de façon que chacun présente une longueur

proche de la moyenne de la longueur du groupe après altération et du groupe qui le précède.

13. Procédé selon la revendication 6, caractérisé en ce que le codage d'essai étant effectué selon une norme de type MPEG à 50 Hertz, avec $N = 12$, $M = 3$ et $Q = 5$, les nombres N et M sont fonction respectivement des coûts moyens de codage des images B et P selon les relations suivantes :

$$(1) N = \text{INT} \left[\frac{389000 - P \cos t}{10000} \right] + 1, \text{ avec } 12 \leq N \leq 30,$$

$$(2) M = \text{INT} \left[\frac{179000 - B \cos t}{20000} \right] + 1, \text{ avec } 1 \leq M \leq 7,$$

10 INT signifiant partie entière.

14. Procédé selon la revendication 13, caractérisé en ce que $1 \leq M \leq 5$.

15 15. Procédé selon la revendication 13 ou 14, caractérisé en ce que, lorsque le coût de codage ($B \cos t$) est supérieur à 179000, le nombre M est déterminé par la relation suivante :

$$(3) M = 5 \cdot \text{INT} \left[\frac{P \cos t}{B \cos t} - 1 \right], \text{ avec } 1 \leq M \leq 7.$$

16. Procédé selon l'une quelconque des revendications 1 à 6, caractérisé en ce qu'on fait varier le nombre M à l'intérieur d'un groupe.

17. Procédé selon l'une quelconque des revendications 2 à 6, caractérisé en ce que la compression est effectuée après le codage d'essai.

18. Procédé selon l'une quelconque des revendications 4 à 6, caractérisé en ce que la compression débute après détermination du paramètre caractérisant la première image B ou la première image P .

19. Procédé selon la revendication 18, caractérisé en ce que la formation d'un groupe codé est interrompue quand son nombre d'images codées est au moins égal au nombre N déterminé à partir de l'image P courante.

20. Dispositif de codage pour la mise en oeuvre du procédé selon la revendication 4, 5 ou 6, caractérisé en ce qu'il comporte un canal pour effectuer le codage d'essai et déterminer

les paramètres M et N et un canal de codage $(78, 74)$ recevant les informations du premier canal pour effectuer le codage proprement dit.

1/2

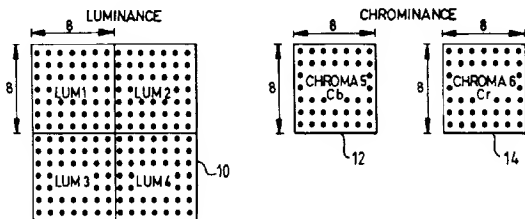


FIG. 1

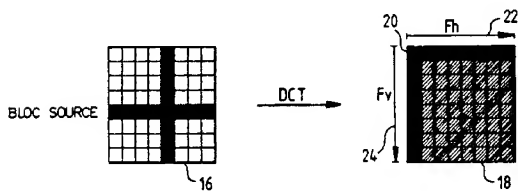


FIG. 2

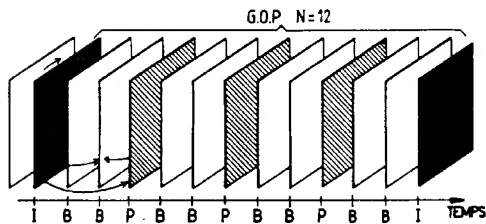


FIG. 3

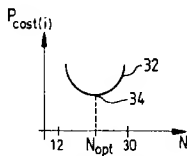


FIG. 4

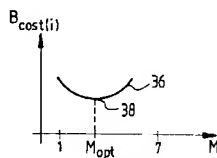


FIG. 5

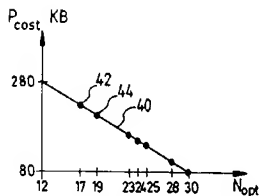


FIG. 6

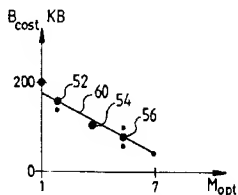


FIG. 7

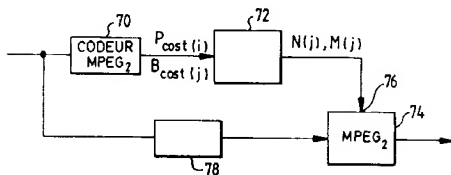


FIG. 8

INSTITUT NATIONAL
de la
PROPRIÉTÉ INDUSTRIELLE

RAPPORT DE RECHERCHE
PRELIMINAIRE

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

N° d'enregistrement
national

FA 562967
FR 9811495

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
A	US 5 694 171 A (KATTO JIRO) 2 décembre 1997 * colonne 5, ligne 39 - colonne 7, ligne 17 * * abrégé; figures 1,6 *	1-20
A	LEE J ET AL: "TEMPORALLY ADAPTIVE MOTION INTERPOLATION EXPLOITING TEMPORAL MASKING IN VISUAL PERCEPTION" IEEE TRANSACTIONS ON IMAGE PROCESSING, vol. 3, no. 5, 1 septembre 1994, pages 513-526, XP000476828 * abrégé *	1-20
A	EP 0 597 647 A (SONY CORP) 18 mai 1994 * revendications 1,5; figure 1 *	1-20
A	EP 0 762 778 A (SONY CORP) 12 mars 1997 * abrégé; figure 5 *	1,11,12
		DOMAINES TECHNIQUES RECHERCHES (Int. Cl. 5)
Date d'achèvement de la recherche		Examinateur
3 juin 1999		La, V
<p>CATÉGORIE DES DOCUMENTS CITÉS</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure D : cité dans la demande I : cité pour d'autres raisons A : membre de la même famille, document correspondant</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général O : divulgation non écrite P : document intermédiaire</p>		

1
EPO FORM 1531 (04/97) (mod. 5)

02 - BR 9904108
04 - CN 1248864
07 - EP 987903
10 - FR 2783388
15 - JP 2000102021
20 - KR 2000023133
27 - ZA 9905802

The method involves coding images according to groups each of which comprises N images, with an I image coded in intra mode, P images predicted as a function of the intra image I or of the preceding P image. Each P image is preceded and followed by n bidirectionally predicted B images, n possibly being zero. The number $M = n + 1$ represents the structure of the group.

At least one parameter characterising the source images which are to be coded according to a group is determined using a test coding (70). The numbers N and M are made to depend on this parameter. In the course of the test coding, defined values are conferred on N, M and on the quantisation interval Q.